NOVEL ACETYLCHOLINESTERASE GENE RESPONSIBLE FOR INSECTICIDE RESISTANCE AND APPLICATIONS THEREOF

The present invention relates to a novel acetylcholinesterase gene responsible for insecticide resistance, in particular in mosquitoes, to the products of this gene (cDNA, protein) and to the applications thereof, in particular for screening novel insecticides and for the genetic detection of resistance to organophosphorus compounds and to carbamates in mosquito populations.

Acetylcholinesterase (AChE, E.C. 3.1.1.7) is an essential enzyme which hydrolyzes acetylcholine in the synapses, thus terminating cholinergic transmissions at neuronal or neuromuscular junctions. The inhibition of AChE prevents the deactivation of the synaptic signal, thus resulting in a loss of control of cholinergic transmission. The biology of acetylcholinesterase has been greatly studied in invertebrates, and in particular insects, since this enzyme is the target for the main classes of pesticides used, organophosphorus compounds and carbamates. However, the massive use of pesticides over the past decades has caused resistant species to emerge. Among the mechanisms of resistance, the selection of mutations making AChE insensitive to insecticides has been observed in many cases (for a review, see Fournier et al., Comp. Biochem. Physiol., 1994, 108, 19-31).

In order to precisely determine the nature of the AChE that is a target for insecticides, and also the mutations responsible for the resistance to the latter, the genes encoding AChEs (*ace* genes) have been isolated in various arthropod (insect and arachnid) species.

The first *ace* gene was identified in drosophila (*Drosophila melanogaster*), by reverse genetics (Hall et al., EMBO J., 1986, 5, 2949-2954). The proof that this gene was involved in insecticide resistance was provided by the demonstration of amino

acid substitutions in the AChE of resistant drosophila, conferring insensitivity to cholinergic insecticides (Mutéro et al., P.N.A.S., 1994, 91, 5922-5926). The studies in *D. melanogaster* therefore appear to indicate the presence of a single *ace* gene in insects, encoding the AChE that is a target for cholinergic insecticides.

However, with the exception of the *ace* gene of another insect, *Musca domestica* (Williamson et al., 1992, in *Multidisciplinary approaches to cholinesterase functions*, Eds Schafferman A. & Velan B., Plenum Press, New-York, pp 83-86; Walsh et al., Biochem. J., 2001, 359, 175-181; Kozaki et al., Insect Biochem. Mol. Biol., 2001, 31, 991-997), the study of the *ace* genes isolated from other insects or else from arachnids, by homology with that of drosophila, indicates that they are not involved in insecticide resistance.

In fact, no mutation in the amino acid sequence of AChE encoded by the *ace* gene of *Aphis gossypii*, of *Nephotettix cincticeps* and of *Boophilus microplus* is observed between resistant and sensitive individuals (Menozzi et al., doctoral thesis from the Paul Sabatier university, Toulouse, 2000; Tomita et al., Insect Biochem. Mol. Biol., 200, 30, 325-333; Baxter et al., Insect Biochem. Mol. Biol., 1998, 28, 581-589; Hernandez et al., J. Med. Entomol., 1999, 36, 764-770), and independent segregation is observed between the *Culex pipiens* and *C. tritaeniorynchus ace* gene and insecticide resistance (Malcolm et al., Insect. Mol. Biol., 1998, 7, 107-120; Mori et al., Insect Mol. Biol., 2001, 10, 197-203).

As regards the other *ace* genes isolated from other insects, their role in insecticide resistance has not been studied (*Lucilia cuprina*: Chen et al., Insect. Biochem. Mol. Biol., 2001, 31, 805-816; *Schizaphis graminum*: Gao et al., Insect. Biochem. Mol. Biol., 2001, 31, 1095-1104) or no insecticide-insensitive form of AChE has been described (*Aedes aegypti, Anopheles gambiae* and *Anopheles stephensi*: Anthony et al., FEBS letters, 1995, 368, 461-465; Malcolm et al., in

Molecular Insect Science, Eds Hageborn et al., Plenum Press, New York, pp 57-65).

Two hypotheses have been put forward to explain the difference in insecticide resistance observed between *Drosophila melanogaster* or *Musca domestica* and the other insects or the arachnids which have been studied: the presence of a "modifier gene" responsible for post-transcriptional or post-translational modifications of AChE, resulting in AChE forms having different catalytic activities, and the presence of a second *ace* gene.

However, no study has made it possible to verify these hypotheses and, consequently, to determine the nature of the gene and that of the target (AChE) involved in insecticide resistance in insects other than *Drosophila melanogaster* and *Musca domestica* or else in arachnids:

- The demonstration, in *C. pipiens*, of two AChE forms having distinct catalytic activities supports the two hypotheses, but the biochemical analysis of these AChEs has not made it possible to determine the nature of the AChE involved in insecticide resistance (Bourguet et al., J. Neurochemistry, 1996, 67, 2115-2123).

- A second *ace* gene has been isolated in arachnids; however, this gene is not involved in insecticide resistance (Hernandez et al., Baxter et al., mentioned above).

- It has not been possible to isolate a second *ace* gene in insects despite many attempts in various species (Menozzi et al., Tomita et al., Mori et al., mentioned above; Severson et al., J. Hered., 1997, 88, 520-524).

It emerges from the above that the nature of the gene and of the target (AChE) involved in the resistance to organophosphorus compounds and to carbamates has not been identified in most insects and in arachnids, in particular in those in which they have been investigated; mention may be made of those which

are the most important in the human and animal health fields and agricultural field, such as pathogen vectors and pests, in particular many mosquitoes such as *Culex pipiens, Aedes aegypti, Anopheles gambiae, Anopheles albimanus* or *Anopheles stephensi*, and crop pests such as *Aphis gossypii, Nephotettix cincticeps* and *Leptinotarsa decemlineata*.

The inventors have identified a novel locus of the *ace* gene in the genome of *Anopheles gambiae* and of 15 different species of mosquitoes, and they have shown that this novel locus, which is not homologous to the locus previously described in *D. melanogaster*, is involved in insecticide resistance in mosquitoes.

This novel gene represents a diagnostic tool for the genetic detection of insecticide (organophosphorus compound, carbamate) resistance in mosquito populations. The AChE encoded by this gene represents a target for the screening of novel molecules that are active on the populations of mosquitoes resistant to the insecticides currently used.

Consequently, a subject of the present invention is a protein, characterized in that it comprises a central catalytic region which has an amino acid sequence selected from the group consisting of the sequence SEQ ID NO. 1 and the sequences exhibiting at least 60% identity or 70% similarity with the sequence SEQ ID NO. 1, with the exclusion of the NCBI sequence AAK0973 corresponding to the *Schizaphis graminum* acetylcholinesterase.

The protein according to the invention represents a novel insect acetylcholinesterase, hereinafter referred to as AchE1, responsible for resistance to organophosphorus compounds and to carbamates, at least in mosquitoes, in particular in *C. pipiens*; the locus encoding said AchE1 is hereinafter referred to as *ace-1*; *ace-2* represents the second *ace* locus, which is not involved in insecticide resistance in mosquitoes. The single *ace* gene present in

Drosophila melanogaster, which is homologous to *ace-2*, is therefore also referred to as *ace-2*.

In accordance with the invention, said central catalytic region contains the catalytic domain of the AChE, and corresponds to that located between positions 70 and 593 of the sequence of *Anopheles gambiae* AChE1 (SEQ ID NO. 3, 643 amino acids); it corresponds to that located, respectively, between positions 100 and 629 of the sequence of *Schizaphis graminum* AChE1 (NCBI AAK0973), 60 and 582 of the sequence of *Culex pipiens* AChE1 (SEQ ID NO. 7), 34 and 593 of the sequence of *Anopheles gambiae* AChE2 (figure 1, SEQ ID NO. 53), and 41 and 601 of the sequence of *Drosophila melanogaster* AChE2 (NCBI AAF54915). This central region which contains the catalytic domain is conserved in vertebrates and invertebrates, whereas the N- and C-terminal ends exhibit great variability between the various species.

In accordance with the invention, the identity of a sequence relative to a reference sequence (SEQ ID NO. 1) is assessed according to the percentage of amino acid residues which are identical, when the sequences corresponding to the catalytic region as defined above are aligned, so as to obtain the maximum correspondence between them.

A protein having an amino acid sequence having at least X% identity with the reference sequence SEQ ID NO. 1 is defined, in the present invention, as a protein which sequence corresponding to the central catalytic region as defined above can include up to 100-X alterations per 100 amino acids of the sequence SEQ ID NO. 1. For the purpose of the present invention, the term "alteration" includes consecutive or dispersed amino acid deletions, substitutions or insertions in the reference sequence. This definition applies, by analogy, to the nucleic acid molecules.

The similarity of a sequence relative to the reference sequence SEQ ID NO. 1 is assessed according to the percentage of amino acid residues which are

identical or which differ by conservative substitutions, when the sequences corresponding to the central catalytic region as defined above are aligned so as to obtain the maximum correspondence between them. For the purpose of the present invention, the term "conservative substitution" is intended to mean the substitution of an amino acid with another which has similar chemical properties (size, charge or polarity), which generally does not modify the functional properties of the protein.

A protein having an amino acid sequence having at least X% similarity with the sequence SEQ ID NO. 1 is defined, in the present invention as a protein which sequence corresponding to the central catalytic region as defined above can include up to 100-X non-conservative alterations per 100 amino acids of the reference sequence. For the purpose of the present invention, the term "non-conservative alterations" includes consecutive or dispersed amino acid deletions, non-conservative substitutions or insertions in the sequence SEQ ID NO. 1.

The comparison of the AChE1 according to the invention with the insect AChEs available on the databases, by alignment of the sequences corresponding to the central region as defined above, using the BLAST program (http://www.ncbi.nlm.nih.gov/gorf/bl2.html, default parameters, inactivated filter) shows that:

- the insect AChE1 and AChE2 sequences exhibit 36-39% identity (53-57% similarity) with one another,
- the insect AChE1 sequences exhibit 65-97% identity (79-98% similarity) with one another,
- the insect AChE2 sequences exhibit 58-99% identity (73-99% similarity) with one another.

In addition, the phylogenetic analysis of the AChEs of the various animal species shows that the AChE1 protein sequences form a significant autonomous group (bootstrap 795/1000), and that the insect AChE1s form a significant distinct subgroup (bootstrap 856/1000).

The AChE1 according to the invention comprises units characteristic of AChEs (figure 1) located at the following positions, respectively, in the sequence SEQ ID NO. 3 and in the reference sequence from *Torpedo marmorata* (SWISSPROT P07962): a canonic unit of the FGESAG type around the serine at position 266 (*200*), which is characteristic of the AChE active site, a choline-binding site (tryptophan residue at position 151 (*84*)), three residues of the catalytic triad (serine, glutamic acid and histidine residues, respectively at positions 266 (*200*), 392 (*327*) and 506 (*440*)), six cysteine residues potentially involved in conserved disulfide bridges ($C_{134(67)}$-$C_{161(94)}$; $C_{320(254)}$-$C_{333(265)}$; $C_{468(402)}$-$C_{589(521)}$), aromatic residues bordering the active site gorge (10 residues) and a phenylalanine residue at position 355 (*290*) but not at position 353 (*288*), which distinguishes invertebrate AChEs from those of vertebrates. It also has a hydrophobic C-terminal peptide corresponding to a glycolipid addition signal, indicating post-translational cleavage of a C-terminal fragment and the addition of a glycolipid anchoring residue as in *Drosophila*; the cysteine residue in the C-terminal sequence preceding the potential site for cleavage of the hydrophobic peptide could be involved in an intermolecular disulfide bond linking the two catalytic subunits of the AChE dimer.

The AChE1 according to the invention differs from the AChE of *Drosophila* (AChE2) by the absence of a hydrophilic insertion of 31 amino acids between the residues located at positions 174 and 175 of the sequence SEQ ID NO. 3 (figure 1); this hydrophilic insertion could be characteristic of AChE2, at least in the Diptera.

The invention encompasses the insect AChE1s sensitive or resistant to organophosphorus compounds and to carbamates.

For the purpose of the present invention, the term "sensitive AChE" is intended to mean an AChE for which the acetylcholinesterase activity is inhibited in

the presence of organophosphorus compounds or of carbamates.

For the purpose of the present invention, the term "resistant AChE" is intended to mean an AChE for which the activity is not inhibited by concentrations of organophosphorus compounds or of carbamates which inhibit 100% of the activity of the corresponding "sensitive AChE" derived from an individual of the same species; this "resistant AChE" differs from the preceding one by the presence of one or more mutations in its amino acid sequence (amino acid substitutions) which modify its sensitivity to acetylcholinesterase inhibitors; among these mutations, mention may be made of the following: F78S, I129V, G227A and F228Y, the amino acids being numbered with reference to the sequence of *Torpedo marmorata* AChE (SWISSPROT P07962).

The acetylcholinesterase activity and the catalytic parameters of the AChEs are measured by conventional enzymatic techniques such as those described in Bourguet et al., mentioned above.

The proteins according to the invention include any natural, synthetic, semi-synthetic or recombinant protein of any prokaryotic or eukaryotic organism, comprising or consisting of an amino acid sequence of an AChE1 protein as defined above. They include in particular the natural proteins isolated from any insect species, and also the recombinant proteins produced in a suitable expression system.

According to an advantageous embodiment of said AChE1, it corresponds to that of an insect which belongs to the order *Diptera*; preferably, said insect is chosen from the family *Culicidae*, from the genera *Culex*, *Aedes* and *Anopheles*.

According to an advantageous arrangement of this embodiment, said AChE1 consists of the sequences SEQ ID NO. 3, SEQ ID NO. 5 and SEQ ID NO. 7, corresponding to the complete sequence of the AChE1 of two strains of *Anopheles gambiae* respectively and of a

strain of *Culex pipiens*, that are sensitive to organophosphorus compounds and to carbamates.

According to another advantageous arrangement of this embodiment, said central catalytic region of the AChE1 corresponds to a sequence selected from the group consisting of the sequences SEQ ID NOs. 8 to 21 representing a fragment of approximately 91 amino acids (fragment K, figure 1), corresponding to that located between positions 445 and 535 of the sequence SEQ ID NO. 3.

A subject of the present invention is also a peptide, characterized in that it consists of a fragment of at least 7 amino acids of the AChE1 protein, as defined above; these fragments are particularly useful for producing antibodies that specifically recognize the AChE1 protein.

A subject of the present invention is also antibodies, characterized in that they are directed against the AChE1 protein or a fragment thereof, as defined above.

In accordance with the invention, said antibodies are either monoclonal antibodies or polyclonal antibodies.

These antibodies may be obtained by conventional methods, known in themselves, comprising in particular the immunization of an animal with a protein or a peptide in accordance with the invention, in order to make it produce antibodies directed against said protein or said peptide.

A subject of the present invention is also an isolated nucleic acid molecule, characterized in that it has a sequence selected from the group consisting of:

- the sequences encoding an AChE1 protein as defined above (cDNA and genomic DNA fragment corresponding to the *ace-1* gene),

- the sequences complementary to the above sequences, which may be sense or antisense, and

- the fragments of at least 8 bp, preferably of 15 bp to 500 bp, of the above sequences.

The invention encompasses the sequences of the alleles of the *ace-1* gene derived from any insect, and also the sequences of the natural (sensitive or resistant alleles) or artificial mutants of the *ace-1* gene encoding a sensitive or resistant AChE1 protein, as defined above.

According to an advantageous embodiment of the invention, said sequence encoding an AChE1 protein is selected from the group consisting of:

a) the sequences SEQ ID NO. 2, SEQ ID NO. 4 and SEQ ID NO. 6 which correspond to the cDNA of the AChE1 protein of amino acid sequence, respectively, SEQ ID NO. 3, SEQ ID NO. 5and SEQ ID NO. 7, , as defined above,

b) the sequences SEQ ID NO. 22 and SEQ ID NO. 23 which correspond to the *ace-1* gene of *Anopheles gambiae* encoding the AChE1 in a), which gene has an exon-intron organization comprising at least 7 exons (table I).

Table I: Intron-exon organization of the *ace-1* gene

	5' site		3' site	
	Position	Sequence	Position	Sequence
Intron 1	nd	nd	1179	ttcag/ACGCA
Intron 2	1315	CTCGG/gtaag	1400	ggcag/ACGCG
Intron 3	1938	CTACG/gtagg	2017	gtcag/CTGGG
Intron 4	2214	CTAAG/gtacg	2301	tccag/AGCAC
Intron 5	3009	ACCGG/gtaag	3075	tacag/CAATC
Intron 6	3248	TACCT/gtaag	3355	aacag/CGAAC

According to another advantageous embodiment of the invention, said fragment is selected from the group consisting of the primers of sequence SEQ ID Nos. 39 to 50 and the fragments of sequences SEQ ID Nos. 24 to 38.

The nucleic acid molecules according to the invention are obtained by conventional methods, known in themselves, according to standard protocols such as those described in *Current Protocols in Molecular Biology (Frederick M. AUSUBEL, 2000, Wiley and son Inc, Library of Congress, USA)*. For example, they can be obtained by amplification of a nucleic acid sequence by PCR or RT-PCR, by screening genomic DNA libraries by hybridization with a homologous probe, or else by total or partial chemical synthesis.

The nucleic acid molecules as defined above can be used as probes or as primers for isolating the *ace-1* gene of other species or alleles of this gene, in particular by screening a genomic DNA or cDNA library, and also for detecting/amplifying nucleic acid molecules (mRNA or genomic DNA) encoding an AChE1 protein as defined above.

These various nucleic acid molecules make it possible to demonstrate the *ace-1* gene, allelic variants of this gene, or a functional alteration of this *ace-1* gene (substantial change in insecticide sensitivity) resulting from a mutation (insertion, deletion or substitution) of one or more nucleotides in said gene.

A subject of the present invention is also a method for detecting insects carrying resistance to insecticides of the organophosphorus compound and carbamate class, characterized in that it comprises:

- preparing a sample of nucleic acids from insects to be tested, and
- detecting, by any suitable means, the presence, in said nucleic acid sample, of a mutation in the *ace-1* gene as defined above.

Said detection is carried out by conventional techniques which are known in themselves, for example: (i) by amplification of a region of said *ace-1* gene liable to contain a mutation, and then detection of said mutation by sequencing, or by digestion with a suitable restriction enzyme, of the PCR product

obtained, or else (ii) by hybridization with a labeled probe specific for a region of said *ace-1* gene liable to contain a mutation, and then direct detection of the mismatches and/or digestion with a suitable restriction enzyme.

Preferably, a fragment of approximately 320 bp (fragment K) is amplified by means of the primers SEQ ID NO. 39 and SEQ ID NO. 40. For example, in mosquitoes, a fragment of sequence SEQ ID NOs. 24 to 38 is obtained, which has mutations between insecticide-sensitive and insecticide-resistant mosquitoes. For example, in *C. pipiens* 3 substitutions are observed in the sequence of the resistant individuals, one of which introduces an *EcoRI* site. Analysis of the restriction profile after PCR amplification of the fragment K and digestion of the products obtained with *EcoRI* (RFLP analysis) makes it possible to rapidly detect the *ace-1* genotype in a population of *C. pipiens*; the presence of a single fragment corresponds to the resistant homozygotes (RR), the presence of 2 fragments of approximately 106 bp and 214 bp corresponds to the sensitive homozygous individuals (SS) and the presence of 3 fragments of 106 bp, 214 bp and 320 bp corresponds to the resistant heterozygous individuals (RS).

A subject of the present invention is also a reagent for detecting insects carrying resistance to organophosphorus compounds and to carbamates, characterized in that it is selected from the group consisting of: the nucleic acid molecules and the fragments thereof as defined above (probes, primers) and the antibodies as defined above.

A subject of the present invention is also a recombinant vector, characterized in that it comprises an insert selected from the group consisting of the nucleic acid molecules encoding an AChE1 protein and the fragments thereof as defined above.

Preferably, said recombinant vector is an expression vector in which said nucleic acid molecule or one of its fragments are placed under the control of

suitable regulatory elements for transcription and for translation.

These vectors are constructed and introduced into host cells by conventional recombinant DNA and genetic engineering methods which are known in themselves. Many vectors into which a nucleic acid molecule of interest may be inserted in order to introduce it into and to maintain it in a eukaryotic or prokaryotic host cell are known in themselves; the choice of a suitable vector depends on the use envisioned for this vector (for example, replication of the sequence of interest, expression of this sequence, maintenance of the sequence in extrachromosomal form or else integration into the host's chromosomal material), and also on the nature of the host cell. For example, viral vectors such as baculoviruses or nonviral vectors such as plasmids may be used. In order to express the AChE1, the *ace-1* cDNA may be placed under the control of a constitutive promoter such as the actin 5C promoter, in a suitable vector, and said recombinant vector is introduced into insect cells such as drosophila cells (Schneider S2 cells).

A subject of the present invention is also prokaryotic or eukaryotic cells modified with a recombinant vector as defined above; preferably, these cells are insect cells.

The recombinant vectors and the modified cells as defined above are useful in particular for producing the AChE1 proteins and peptides according to the invention.

A subject of the present invention is also a transgenic invertebrate animal, characterized in that it contains cells modified with at least one nucleic acid molecule as defined above; preferably, said animal is an insect.

The transgenic animals and the modified cells as defined above are useful in particular for screening insecticidal substances.

A subject of the present invention is also a method for screening an insecticidal substance, characterized in that it comprises:

a) bringing the test substance into contact with an AChE1 protein selected from: the AChE1 protein isolated according to the invention, or an extract of modified cells or a biological sample from a transgenic animal containing said AChE1 protein, as defined above, in the presence of acetylcholine or of one of its derivatives,

b) measuring, by any suitable means, the acetylcholinesterase activity of the mixture obtained in a), and

c) selecting the substances capable of inhibiting said activity.

A subject of the present invention is also a method for screening an insecticidal substance, characterized in that it comprises:

- bringing a transgenic animal as defined above into contact with the test substance, and
- measuring the animal's survival.

Advantageously, said screening methods use AChE1s resistant to organophosphorus compounds or to carbamates, or else cells or transgenic animals containing them.

A subject of the present invention is also a reagent for screening insecticidal substances, characterized in that it is selected from the group consisting of the AChE1 proteins, the recombinant vectors, the modified cells and the transgenic animals as defined above.

Insecticidal substances capable of inhibiting the acetylcholinesterase activity of the AChE1 proteins resistant to insecticides of the organophosphorus compound and carbamate class commonly used have applications: in animal and human health, for controlling pathogen vectors (for example *Aedes aegypti*, a vector of arboviroses such as dengue and yellow fever, *Culex pipiens*, a West-Nile virus vector, *Anopheles gambiae*,

an African vector of the agent for malaria, etc.) and in the agricultural field, for controlling insect pests which devastate harvests (for example the Colorado potato beetle (*Leptinotarsa decemlineata*) which attacks potatoes, aphid pests such as *Aphis gossypii* and *Myzus persicae*, etc.).

A subject of the invention is also a detection and/or screening kit for carrying out the methods as defined above, characterized in that it includes at least one reagent as defined above.

Besides the above provisions, the invention also comprises other provisions which will emerge from the following description, which refers to examples of implementation of the *ace-1* gene and of its products (cDNA, protein) according to the present invention and also to the table summarizing the sequences of the application and to the attached drawings in which:

- figure 1 illustrates the amino acid sequence alignment for the AChE1 proteins of *Anopheles gambiae, Schizaphis graminum, An. stephensi, Aedes aegypti, Drosophila melanogaster, Lucilia cuprina, Musca domestica* and *Culex pipiens*. By convention, the amino acids are numbered with reference to the AChE sequence from the torpedo fish (*Torpedo marmorata*; SWISSPROT P07962). The N- and C-terminal sequences are not represented due to their variability. The amino acids that are conserved between AChE1 and AChE2 are indicated in gray. The amino acids specific for AChE2 are indicated in black. The 3 residues representing the catalytic triad (S_{200}, E_{327} and H_{440}) have boxes drawn round them. The choline-binding site (W_{84}) is underlined. The circles represent the position of the 14 aromatic residues bordering the active site gorge in the *Torpedo* AChE, 10 of which are present in all the AChE1s and AChE2s (solid circles), the others not being conserved (open circles). Three intramolecular disulfide bonds between cysteine residues are indicated. The horizontal arrow indicates the position of the fragment K (amplified by means of the primers PdirAGSG and

PrevAGSG). The hypervariable region of AChE2, which is absent in AChE1, is enclosed in a box;

- figure 2 illustrates the genetic detection of mosquitoes resistant to organophosphorus compounds and to carbamates by PCR-RFLP:

- figure 2A represents the comparison of the amino acid sequence of fragment K of various mosquito species: Cx Pip (*Culex pipiens*), Ae alb (*Aedes albopicus*), Ae aeg (*Aedes aegypti*), An alb (*Anopheles albimanus*), An gamb (*Anopheles gambiae*), An fun (*Anopheles funestus*), An nil (*Anopheles nili*), An sac (*Anopheles sacharovi*), An pse (*Anopheles pseudopunctipennis*). The variant amino acids are shaded. The following sequences are identical: *An. darlingi* and *An. albimanus*; *An. sundaicus*, *An. gambiae* and *An. arbiensis*; *An. moucheti*, *An. funestus* and *An. minimus*; *An. stephensi* and *An. saccharovi*;

- figure 2B illustrates the comparison of the nucleotide sequences corresponding to fragment K of the sensitive (S-LAB) and resistant (SR) strains. The variant nucleotides are shaded (t → c at position 3; a → g at position 84: the *EcoRI* site (gaattc) located around this position, used for the PCR-RFLP analysis, is present only in the S-LAB strain; c → t at position 173). Figure 2C illustrates the restriction profiles obtained after agarose gel electrophoresis of the products of digestion with *EcoRI*, of fragment K amplified by PCR. The sensitive homozygous S-LAB strain has a profile characterized by 2 bands (214 bp and 106 bp), the resistant homozygous strain has a profile characterized by a single band of 320 bp and the resistant mosquitoes derived from the back cross have a heterozygous profile characterized by 3 bands (320 bp, 214 bp and 106 bp);

- figure 3 illustrates the phylogenetic tree for the AChE proteins. The phylogenetic analysis was carried out using 47 sequences of AChE proteins from 35 different species originating from the ESTHER database (http://www.ensam.inra.fr/cgi-bin/ace/index). The

sequences were aligned and a tree was constructed as described in example 1. Only the nodes corresponding to "bootstrap" values > 50% (i.e. scores greater than 500) are indicated. The scale represents a divergence of 10%. Agam: *An. gambiae*; Aeg: *Aedes aegypti*; Aste: *Anopheles stephensi*; Cpip: *Culex pipiens*; Dme1: *Drosophila melanogaster*; Lcup: *Lucilia cuprina*; Mdom: *Musca domestica*; Ldec: *Leptinotarsa decemlineata*; Ame1: *Apis mellifera*; Ncin: *Naphotettix cincticeps*; Sgra: *Schizaphis graminum*; Rapp: *Rhipicephalus appendiculatus*; Bmic: *Boophilus microplus*; Bdec: *Boophilus decoloratus*; Hsap: *Homo sapiens*; Btau: *Bos taurus*; Fcat: *Felix catus*; Ocun: *Oryctolagus cuniculus*; Rnor: *Rattus norvegicus*; Mmus: *Mus musculus*; Ggal: *Gallus gallus*; Drer: *Danio reno*; Eele: *Electrophorus electricus*; Tamr: *Torpedo marmorata*; Tcal: *Torpedo californica*; Bfas: *Bungarus fasciatus*; Mglu: *Myxine glutinosa*; Bflo: *Branchiostoma floridae*; Blan: *Branchiostoma lanceolatum*; Cint: *Ciona intestinalis*; Csav: *Ciona savignyi*; Cele: *Caenorhabditis elegans*; Cbrig: *Caenorhabditis briggsae*; Dviv: *Dictyocaulus viviparus*; Lopa: *Loligo opalescens*;

- figure 4 illustrates the cladogram for the AChE1 and AChE2 proteins. The sequences of the AChE1 and AChE2 proteins were processed as in figure 1. The Bmic sequence was added as external sequence in order to define the origin of the tree. The boxes marked with an asterisk represent the proteins encoded by a gene which segregates with insecticide resistance. The open boxes represent the proteins encoded by a gene which does not segregate with insecticide resistance. The scale corresponds to a divergence of 10%;

Table II: Sequence listing

Identification number	Sequence
SEQ ID No.: 1	Fragment of the central region of the *Anopheles gambiae* AChE1 protein (positions 70 to 593 of SEQ ID No. 3).
SEQ ID No.: 2	*Anopheles gambiae* AChE1 cDNA
SEQ ID No.: 3	*Anopheles gambiae* AChE1 protein
SEQ ID No.: 4	*Anopheles gambiae* (strain KISUMU) AChE1 cDNA

SEQ ID No.: 5	*Anopheles gambiae* (strain KISUMU) AChE1 protein
SEQ ID No.: 6	*Culux pipiens* strain S-LAB AChE1 cDNA
SEQ ID No.: 7	*Culex pipiens* strain S-LAB AChE1 protein
SEQ ID No.: 8	Peptide fragment K AChE1 *Culex pipiens*
SEQ ID No.: 9	Peptide fragment K AChE1 *Aedes aegypti*
SEQ ID No.: 10	Peptide fragment K AChE1 *Aedes albopictus*
SEQ ID No.: 11	Peptide fragment K peptide AChE1 *Anopheles darlingi*
SEQ ID No.: 12	Peptide fragment K AChE1 *An. sundaicus*
SEQ ID No.: 13	Peptide fragment K AChE1 *An. minimus*
SEQ ID No.: 14	Peptide fragment K AChE1 *An. moucheti*
SEQ ID No.: 15	Peptide fragment K AChE1 *An. arabiensis*
SEQ ID No.: 16	Peptide fragment K AChE1 *An. funestus*
SEQ ID No.: 17	Peptide fragment K AChE1 *An. pseudopunctipennis*
SEQ ID No.: 18	Peptide fragment K AChE1 *An. sacharovi*
SEQ ID No.: 19	Peptide fragment K AChE1 *An. stephensi*
SEQ ID No.: 20	Peptide fragment K AChE1 *An. albimanus*
SEQ ID No.: 21	Peptide fragment K AChE1 *An. nili*
SEQ ID No.: 22	*An. gambiae ace-1* gene
SEQ ID No.: 23	*An. gambiae* KISUMU *ace-1* gene
SEQ ID No.: 24	Nucleotide fragment K AChE1 *C. pipiens* (strain S-LAB)
SEQ ID No.: 25	Nucleotide fragment K AChE1 *C. pipiens* (strain SR)
SEQ ID No.: 26	Nucleotide fragment K AChE1 *Aedes aegypti*
SEQ ID No.: 27	Nucleotide fragment K AChE1 *Aedes albopictus*
SEQ ID No.: 28	Nucleotide fragment K AChE1 *Anopheles darlingi*
SEQ ID No.: 29	Nucleotide fragment K AChE1 *An. sundaicus*
SEQ ID No.: 30	Nucleotide fragment K AChE1 *An. minimus*
SEQ ID No.: 31	Nucleotide fragment K AChE1 *An. moucheti*
SEQ ID No.: 32	Nucleotide fragment K AChE1 *An. arabiensis*
SEQ ID No.: 33	Nucleotide fragment K AChE1 *An. funestus*
SEQ ID No.: 34	Nucleotide fragment K AChE1 *An. pseudopunctipennis* (
SEQ ID No.: 35	Nucleotide fragment K AChE1 *An. sacharovi*
SEQ ID No.: 36	Nucleotide fragment K AChE1 *An. stephensi*
SEQ ID No.: 37	Nucleotide fragment K AChE1 *An. albimanus*
SEQ ID No.: 38	Nucleotide fragment K AChE1 *An. nili*
SEQ ID No.: 39	Primer PkdirAGSG
SEQ ID No.: 40	Primer PkrevAGSG
SEQ ID No.: 41	Primer PbdirAGSG
SEQ ID No.: 42	Primer PbrevAGSG
SEQ ID No.: 43	Primer culex-bdir1
SEQ ID No.: 44	Primer culex-krev1
SEQ ID No.: 45	Primer AG1-Adir
SEQ ID No.: 46	Primer AG1-Arev
SEQ ID No.: 47	Primer AG1-Bdir
SEQ ID No.: 48	Primer AG1-Brev
SEQ ID No.: 49	Primer AG1-Cdir

SEQ ID No.: 50	Primer AG1-Crev
SEQ ID No.: 51	*Ciona intestinalis* AChE1 protein
SEQ ID No.: 52	*Ciona savignyi* AChE1 protein
SEQ ID No.: 53	*Anopheles gambiae* AChE2 protein

EXAMPLE 1: Materials and methods

a) Strains and crosses

Five *C. pipiens* strains were used: S-LAB, a standard insecticide-sensitive strain (Georghiou et al., 1996, Bull. Wld. Hlth Org., 35, 691-708), SA1, SA4 and EDIT, which have a single insecticide-sensitive AChE, and SR which is homozygous for an insecticide-insensitive AChE (Berticat et al., Genet. Res., 2002, 79, 41-47). The strains having a sensitive and insensitive AChE are referred to, respectively, as S and R.

b) *ace* gene nomenclature and amino acid sequence numbering

ace-1 represents the locus encoding a cholinergic AChE responsible for resistance to organophosphorus compounds and to carbamates in *C. pipiens* (AChE1), previously called *Ace.1* (Raymond et al., Genetica, 2001, 112/113, 287-296). *ace-2* represents the second *ace* locus, which is not involved in insecticide resistance in *C. pipiens* (previously called *Ace.2*), the function of which is unknown in *C. pipiens*. The single *ace* gene present in *Drosophila melanogaster*, which is homologous to *ace-2*, is therefore similarly named.

In the analyses which follow, the positions of the amino acid residues are indicated with reference to the sequence of the torpedo fish AChE [*Torpedo marmorata*; GENBANK P04058], according to the nomenclature recommended by Massoulié et al., 1992, *In Multidisciplinary approaches to cholinesterase functions, eds, Schafferman, A. & Velan, B. (Plenum Press New York), p 285-288*].

c) Analysis of transmission of the *ace-1* gene

With the females being indicated first, F1 crosses (S X R) and back crosses (F1 X S-LAB and S-LAB X F1) were obtained by mass crossing of adults. A few larvae derived from the back crosses were treated with a dose of carbamate (propoxur, 4 mg/l) which kills 100% of sensitive larvae. The linkage between *ace-1* and propoxur resistance was studied by RFLP in the surviving larvae, based on a 320 bp PCR product making it possible to identify the S and R alleles. The experiments were carried out independently, with S = SA1, S = SA4 and S = EDIT.

d) Sequence analysis and gene assembly

All the sequence analyses were carried out based on the crude sequences of *Anopheles gambiae* available on the INFOBIOGEN server (http://www.infobiogen.fr) and the tools available on the site (http://www.ncbi.nlm.nih.gov/blast/blast). The genomic sequences encoding an AChE were identified using the TBLASTN and BLAST programs (Altschul et al., J. Biol. Mol., 1990, 215, 403-410). The genomic sequences identified were assembled using the ABI Prism Auto-Assembler program (v2.1, Perkin Elmer). The sequences were verified and corrected using the Ensembl Trace Server program (http://trace.ensembl.org/). Two concatenations of, respectively, 5195 and 6975 base pairs, encoding respectively AChE1 and AChE2, were assembled from, respectively, 64 and 74 independent sequences (mean redundancy of 10.5 and 6.5). The exons and the protein sequences were identified using a combination between the FGENESH (http://www.sanger.uk) and BLASTX (http://www.ncbi.nlm.nih.gov) programs. The genomic sequences of ascidian AChE were assembled from crude sequences deposited in the databases of the NCBI (*Ciona savignyi*) and of the Doe Joint Institute (*Ciona intestinalis*, http://www.jgi.doe.gov/programs/ciona/ciona-mainage.html). The searches in the *Drosophila* databases were carried out using Flybase (http://www.fruitfly.org/).

e) Sequence comparisons

The sequences of the *Anopheles gambiae* AChE1 and AChE2 proteins deduced from the genomic sequences and the peptide sequences deduced from PCR fragments of *C. pipiens* and *A. aegypti* were aligned with those of known AChEs, by means of the ClustalW program, using a BLOSUM matrix and default parameters (Thompson et al., N.A.R., 1994, 22, 4673-4680).

A phylogenetic tree was constructed using the neighbor-joining algorithm of version DDBJ of Clustal W (http://hypernig.nig.ac.jp/homology/ex-clustalw-e.shtml). *Bootstrap* analysis (1000 counts and 111 entry values) was used to evaluate the degrees of confidence for the topology of the tree. The construction of the trees was carried out using the Treeview program (v1.6.6).

f) Accession numbers

The numbers of the sequences (accession numbers in the databases or the identifying numbers in the sequence listing) which were used for the genetic analysis are as follows:

- Craniata: *Homo sapiens*: NP_00046; *Bos taurus*: P23795; *Felix catus*: O6763; *Oryctolagus cuniculus*: Q29499; *Rattus norvegicus*: P36136; *Mus musculus*: P21836; *Gallus gallus*: CAC37792; *Danio reno*: Q9DDE3; *Electrophorus electricus*: 6730113; *Torpedo marmorata*: P07692; *Torpedo californica*: P04058; *Bungarus fasciatus*: Q92035; *Myxine glutinosa*: Q92081.

- Cephalocordes: *Branchiostoma floridae*: O76998 and 076999; *Branchiostoma lanceolatum*: Q95000 and Q95001.

- Urocordes: *Ciona intestinalis*: SEQ ID NO 51; *Ciona savignyi*: SEQ ID NO 52.

- Nematodes: *Caenorhabditis elegans* (1 to 4): P38433, O61371, O61459 and 061372; *Caenorhabditis briggsae* (1 to 4) Q27459, O61378Q9NDG9 and Q9NDG8; *Dictyocaulus viviparus*: Q9GPLO.

- Insects: *Anopheles gambiae* (1 and 2): SEQ ID NO 3 and SEQ ID NO 53; *Aedes aegypti* (1 and 2): SEQ ID NO 9 and AAB3500; *An. stephensi*: P56161; *Culex pipiens*:

SEQ ID NO 7 (*ace-1*) and Esther database for *ace-2*; *Drosophila melanogaster*: P07140; *Lucilia cuprina*: P91954; *Musca domestica*: AAK69132.1; *Leptinotarsa decemlineata*: Q27677; *Apis mellifera*: AAG43568; *Nephotettix cincticeps*: AF145235_1; *Schizaphis graminum*: Q9BMJ1.

- Arachnids: *Rhipicephalus appendiculatus*: O62563; *Boophilus microplus* (1 and 2): O45210 and O61864; *Boophilus decoloratus*: O61987.

- Mollusks: *Loligo opalescens*: O97110.

g) Cloning of the fragment K and genotyping of *ace-1* in *Culex pipiens*

The mosquito DNA was extracted as described in Rogers et al. [*Plant Molecular Biology manual, 1988, eds. Gelvin, S.B.1 Schilperoot, R.A. (Kluwer Academic Publishers, Boston), Vol. A6, p 1-10*]. The oligonucleotides PkdirAGSG (5'-ATMGWGTTYGAGTACACSGAYTGG-3', SEQ ID NO 39) and Pkrev AGSG (5'-GGCAAARTTKGWCCAGTATCKCAT-3', SEQ ID NO 40) amplify a 320 bp fragment (fragment K) from the genomic DNA of several mosquitoes. 30 PCR amplification cycles were carried out under the following conditions: 94°C for 30 s, 50°C for 30 s and 72°C for 30 s. The sequences were determined directly on the PCR products on an ABI prism 310 sequencer, using the *Big Dye Terminator* kit.

The genotyping of *ace-1 Culex* was carried out under the following conditions: the fragments K obtained as described above were digested with *EcoRI* and the digestion product was separated by electrophoresis on a 2% agarose gel. The restriction profiles show: 1 band (320 bp) in the resistant homozygous RR mosquitoes, 2 bands (106 bp and 214 bp) in the homozygous SS mosquitoes and 3 bands (103 bp, 214 bp and 320 bp) in the heterozygous RS mosquitoes.

h) Cloning of the *ace-1* cDNA in sensitive and resistant individuals

The cDNA of the *Culex pipiens ace-1* gene was obtained from the RNA extracted from individuals of the reference sensitive strain S-LAB, at the very first

larval stage of the development L1. The reverse transcription was carried out with an 18T oligonucleotide and SuperScriptIIRNaseH (INVITROGEN), according to the conditions recommended by the manufacturer.

Two cDNA fragments were amplified by PCR using degenerate oligonucleotides obtained from the alignment of the sequences of the *Anopheles gambiae* and *Schizaphis graminum ace-1* genes:

- a fragment b (193 bp) was amplified using the pair of primers PbdirAGSG (5'GGYGCKACMATGTGGAAYCC3', SEQ ID NO 41) and PbrevAGSG (5'ACCAMRATCACGTTYTCYTCCGAC3', SEQ ID NO 42);
- a fragment k (320 bp) was amplified using the pair of primers PkdirAGSG (5'ATMGWGTTYGAGTACACSGAYTGG3', SEQ ID NO 39) and PkrevAGSG (5'GGCAAARTTKGWCCAGTATCKCAT3', SEQ ID NO 40).

The fragments b and k thus obtained were then cloned and sequenced, according to conventional techniques known in themselves to those skilled in the art, as described in Current Protocols in Molecular Biology (*Frederick M. AUSUBEL, 2000, Wiley and Son Inc, Library of Congress, USA*).

A larger cDNA fragment was amplified by PCR, using the *Culex pipiens*-specific primers deduced from the sequences of the fragments b and k obtained above. Namely:

- a fragment CulexA (1127 bp) was amplified by PCR using the pair of primers: culex-bdir1 (5'TACATCAACGTGGTCGTGCCACG3', SEQ ID NO 43) and culex-krev1 (5'GTCACGGTTGCTGTTCGGG3', SEQ ID NO 44). The 1127 bp fragment CulexA thus obtained was then cloned and sequenced, as above.

The ends of the cDNAs were amplified by the RACE (*Rapid Amplification of cDNA Ends*) technique using a commercial kit (*Gene Racer* kit (INVITROGEN)) according to the conditions indicated in the instruction booklet. They were subsequently cloned and then sequenced, as above.

i) Cloning of the *ace-1* gene in sensitive and resistant individuals

The sequence of the genomic DNA of the strain *A. gambiae* KISUMU (reference sensitive strain from West Africa) and of the strain YAO (resistant strain isolated in Ivory Coast) has been obtained from homozygous mosquitoes.

More precisely, the DNA of mosquitoes has been extracted as described in Rogers et al. [*Plant Molecular Biology manual, 1988, eds. Gelvin, S.I.1 Schilperoot, R.A. (Kluwer Academic Publishers, Boston) Vol. A6, p 1-10*].

3 overlapping fragments (A, B and C) were amplified under the following conditions: 94°C for 30 s, 50°C for 30 s and 72°C for 30 s (30 cycles), using the primers synthesized from the sequence of the *ace-1* gene. Namely:

- the fragment A (1130 bp) was amplified using the pair of primers AG1-Adir (5'CGACGCCACCTTCACA3', SEQ ID NO 45) and AG1-Arev (5'GATGGCCCGCTGGAACAGAT3', SEQ ID NO 46),

- the fragment B (1167 bp) was amplified using the pair of primers AG1-Bdir (5'GGGTGCGGGACAACATTCAC3', SEQ ID NO 47) and AG1-brev (5'CCCCGACCGACGAAGGA3', SEQ ID NO 48), and

- the fragment C (876 bp) was amplified using the pair of primers AG1-Cdir (5'AGATGGTGGGCGACTATCAC3' SEQ ID NO 49) and AG1-Crev (5'CTCGTCCGCCACCACTTGTT3', SEQ ID NO 50).

The sequences of the fragments A, B and C were determined directly on the PCR products, by means of internal oligonucleotides, included in these fragments, using the *Big Dye Terminator* kit and an ABI prism 310 sequencer.

EXAMPLE 2: Demonstration of 2 *ace* genes in *Anopheles gambiae*

Genes homologous to the human and drosophila acetylcholinesterase genes were sought based on sequence fragments deposited in the databases, using

the TBLASTN program. Two groups of distinct fragments encoding an AChE very similar to that of drosophila were identified. Two genes of, respectively, 6975 bp (*ace-1*) and 5195 bp (*ace-2*) were reconstructed from overlapping fragments of each group. Analysis of the genes using the FGENESH and BLASTX programs shows that the *ace-1* and *ace-2* genes consist, respectively, of at least 7 and 8 exons encoding proteins of approximately 534 and 569 amino acids, respectively referred to as AChE1 and AChE2. However, this analysis did not make it possible to determine with certainty the sequence of the 5' and 3' ends of the cDNA and the NH_2 and COOH sequences of the corresponding proteins, which are not conserved between the various AChEs.

The amino acid sequence analysis confirms that the AChE1 and AChE2 proteins are highly homologous to the AChE of *Drosophila* (BLASTP : $P < e^{-180}$) and contain a canonic FGESAG motif around tne serine at position 200, with reference to the sequence of the *Torpedo* AChE (S_{200}, figure 1), which is characteristic of the AChE active site. In addition, other motifs characteristic of AChEs were also found in the two sequences (AChE1 and AChE2): the choline-binding site (tryptophan residue at position 84, W84), the three residues of the catalytic triad (serine, glutamic acid and histidine residues, respectively at positions 200, 327 and 440: S_{200}, E_{327} and H_{440}), the six cysteine residues potentially involved in conserved disulfide bridges (C_{67}-C_{94}; C_{254}-C_{265}; C_{402}-C_{521}), and aromatic residues bordering the active site gorge (10 and 11 residues, respectively, for AChE1 and AChE2).

In the two sequences, the presence of a phenylalanine residue is observed at position 290 (F290), but not at position 288; this characteristic common to invertebrate AChEs is responsible for a broader substrate specificity of invertebrate AChEs, compared with those of vertebrates.

Analysis of the C-terminal sequences of Diptera AChEs shows the presence of a hydrophobic peptide

corresponding to a signal for the addition of a glycolipid, indicating post-translational cleavage of a C-terminal fragment and the addition of a glycolipid anchoring residue as in *Drosophila*, and other mosquito species. In all the sequences, the presence of a cysteine residue is also observed in the C-terminal sequence preceding the potential site of cleavage of the hydrophobic peptide. This cysteine residue could be involved in an intermolecular disulfide bond, linking the two catalytic subunits of the AChE dimer.

The AChE1 and AChE2 proteins of *An. gambiae* exhibit 53% similarity between one another and show, respectively: 76% and 55% similarity with the AChE of *Schizaphis graminum* (NCBI accession number AAK09373 or GENBANK accession number 12958609), 53% and 98% similarity with the AChE of *An. stephensi* (GENBANK 2494391), 54% and 95% similarity with the AChE of *Aedes aegypti* (GENBANK 2133626), and 52% and 83% similarity with the AChE of *Drosophila* (GENBANK 17136862).

The major difference between AChE1 and AChE2 lies in an insertion of 31 amino acids in the sequence of AChE2 (figure 1). This sequence, referred to as "hydrophilic insertion" in the AChE of *Drosophila* is absent in vertebrate and nematode AChEs and could be characteristic of AChE2, at least in the diptera.

These results demonstrate the presence of two *ace* genes in the genome of *Anopheles gambiae*, one encoding AChE1 which is related to the AChE of *Schizaphis graminum*, and the other encoding AChE2 which is related to the AChE of *Drosophila* and to the known AChEs of mosquitoes. The presence of other *ace* genes in *An. gambiae* is highly improbable insofar as complementary searches in the databases of the *An. gambiae* genome, using less stringent parameters, detected only sequences encoding alpha-esterases (EC 3.1.1) and carboxylesterases (EC 3.1.1.1).

EXAMPLE 3: Demonstration of a single *ace* gene in *Drosophila melanogaster*

The presence of a gene homologous to the *ace-1* gene was sought in the genome of *Drosophila*. TBLASTN searches made it possible to detect the *ace* gene identified above, homologous to the *ace-2* gene of *Anopheles gambiae*, but did not make it possible to detect other sequences homologous to the *ace-1* gene. Searches using less stringent parameters made it possible to detect only alpha- and carboxylesterases. These results demonstrate that the drosophila genome contains a single *ace* gene (*ace-2*).

EXAMPLE 4: Demonstration of at least two *ace* genes in the other mosquito species

The presence of the *ace-1* gene in the genome of other mosquito species was analyzed by PCR using degenerate oligonucleotides (PdirAGSG and PrevAGSG, SEQ ID NOs. 39 and 40) for amplifying an exon fragment (fragment K of approximately 320 bp, figure 1), corresponding to sequences that are conserved between the AChE1 sequences of *An. gambiae* and *Schizaphis graminum*, but divergent between the AChE1 and AChE2 sequences of *An. gambiae*.

The sequence of the PCR products obtained from the genomic DNA of various mosquito species shows a very high percentage identity between the sequences of *Anopheles, Culex* and *Aedes*. In addition, most of the substitutions are silent since the amino acid sequences deduced from these nucleotide sequences only differ from one another by 5 to 6 amino acids (figure 2A). The fragment K was also amplified by RT-PCR from the mRNA of *C. pipiens*, indicating that the *ace-1* gene is expressed in the form of mRNA; this result is in agreement with the existence, in *C. pipiens*, of two AChEs having distinct catalytic properties.

EXAMPLE 5: Analysis of the linkage between the *ace-1* gene and insecticide resistance

In order to analyze the linkage between the *ace-1* gene and insecticide resistance, the fragment K

amplified from the genomic DNA of resistant *C. pipiens* (strain R) was sequenced. Comparison of the fragment K sequences between the S and R strains shows differences of 3 nucleotides (silent substitutions, figure 2B). One of these substitutions affects an *EcoRI* site, which makes it possible to readily differentiate the *ace-1* loci of the S and R strains by PCR-RFLP: the restriction profiles show 1 band (320 bp) in the resistant homozygous individuals, 2 bands (106 bp and 214 bp) in the homozygous SS mosquitoes and 3 bands (103 bp, 214 bp and 320 bp) in the heterozygous RS mosquitoes (figure 2C).

The linkage between the *ace-1* gene and propoxur resistance was studied, in triplicate, in the following way: larvae from a backcross (S x R) x S were treated for a dose that is lethal for sensitive individuals and the *ace-1* genotype was analyzed in the survivors, by PCR-RFLP.

The results show that exposure to propoxur kills 50% of the larvae in all the back crosses, i.e. all the sensitive individuals. All the surviving larvae (100 for each back cross, 300 in total) show a heterozygous profile by RFLP, indicating that they all have a copy of the *ace-1* gene of the R strain.

These results demonstrate that the *ace-1* gene is very closely linked with the insecticide resistance (less than 1% recombination with a degree of confidence of 0.05).

EXAMPLE 6: Analysis of the phylogeny of the *ace-1* and *ace-2* genes

Phylogenetic trees were constructed from the sequences of the conserved regions of *An. gambiae* AChEs (SEQ ID NO 1 and fragment 34-393 of the sequence SEQ ID NO 53, figure 1), of the fragments K of *C. pipiens* and *Aedes aegypti* (SEQ ID NOs. 8 and 9) and from 33 AChE sequences available in GENBANK, using the neighbor-joining method, as described in the materials and methods.

Figure 3 illustrates the heterogeneity of the number of *ace* genes in the course of the evolution of the animal kingdom. Among the chordates, the cephalochordates have at least two *ace* genes, whereas the urochordates have only a single one, as deduced from the analysis of their genome. Among the arthropods, the Diptera have either a single *ace* gene (*Drosophila* of the suborder Brachycera) or two *ace* genes (mosquitoes of the suborder nematocera). The topology of the tree shows that these two *ace* genes became duplicated very early in the course of evolution, probably before separation between protostomes and deuterostomes. These results are supported by the fact that mollusk, nematode and arthropod AChEs branch from the sequences of chordata (*craniatia*, cephalochordata and urochordata). The results show that arthropods and nematodes have a related AChE.

These results indicate that the *ace-1* and *ace-2* genes identified in insects originate from a very distant duplication event and that the absence of the *ace-1* gene, at least in certain species of the suborder Brachycera (*Drosophila*), results from the loss of an *ace* gene rather than from a recent duplication of the *ace* gene in the nematocera. These results also suggest that the extrapolations made from studies in *D. melanogaster* are to be considered with reservation insofar as the situation in *Drosophila* is representative neither of the Diptera nor of the entire insect class.

EXAMPLE 7: Determination of the complete sequenceof the *ace-1* cDNA

The *ace-1* cDNA was cloned from a strain of *Anopheles gambiae* of West Africa (strain KISUMU) and from a strain of *Culex pipiens* (strain S-LAB), as described in the materials and methods.

The complete sequence of the cDNA corresponds, respectively, to the sequences SEQ ID NO 4 and SEQ ID NO 6 which encode respectively a 643 and 702 amino acid protein (SEQ ID NO 5 and SEQ ID NO 7).

EXAMPLE 8: Determination of the complete sequence of the *ace-1* gene

The complete sequence of the *ace-1* gene was determined from the genomic DNA of a strain of *Anopheles gambiae* from West Africa (strain KISUMU), as described in the materials and methods.

The complete sequence of the *ace-1* gene corresponds to the sequence SEQ ID NO 23 which has an intron-exon organization comprising at least 7 exons (table I).

As it emerges from the above, the invention is in no way limited to its methods of implementation, execution and application which have just been described more explicitly; on the contrary, it encompasses all the variants thereof which may occur to those skilled in the art, without departing from the context or the scope of the present invention.

CLAIMS

1. An insect acetylcholinesterase, characterized in that it comprises a central catalytic region which has an amino acid sequence selected from the group consisting of the sequence SEQ ID NO 1 and the sequences exhibiting at least 60% identity or 70% similarity with the sequence SEQ ID NO 1, with the exclusion of the acetylcholinesterase of NCBI sequence AAK0973.

2. The insect acetylcholinesterase as claimed in claim 1, characterized in that it corresponds to that of an insect of the family *Culicidae*, chosen from the genera *Culex, Aedes* and *Anopheles*.

3. The acetylcholinesterase as claimed in claim 2, characterized in that it consists in a sequence selected from the group consisting of the sequences SEQ ID NO 3, SEQ ID NO 5 and SEQ ID NO 7.

4. The acetylcholinesterase as claimed in claim 2, characterized in that said central catalytic region comprises a sequence selected from the group consisting of the sequences SEQ ID NOs. 8 to 21.

5. A peptide, characterized in that it consists of a fragment of at least 7 amino acids of the acetylcholinesterase as claimed in any one of claims 1 to 4.

9. An isolated nucleic acid molecule, characterized in that it has a sequence selected from the group consisting of:

- the sequences encoding an acetylcholinesterase as claimed in any one of claims 1 to 4 (cDNA and *ace-1* gene),
- the sequences complementary to the above sequences, which may be sense or antisense, and
- the fragments of at least 8 bp, preferably of 15 bp to 500 bp, of the above sequences.

7. The nucleic acid molecule as claimed in claim 6, characterized in that it is selected from the group consisting of the cDNA of sequences SEQ ID NO 2,

SEQ ID NO 4 and SEQ ID NO 6 and the genomic DNAs of sequences SEQ ID NO 22 and SEQ ID NO 23.

8. The nucleic acid molecule as claimed in claim 6, characterized in that it is selected from the group consisting of the primers of sequence SEQ ID NOs. 39 to 50 and the fragments of sequences SEQ ID NOs. 24 to 38.

9. A method for detecting insects carrying resistance to insecticides of the organophosphorus compound and carbamate class, characterized in that it comprises:

- preparing a sample of nucleic acids from insects to be tested, and

- detecting, by any suitable means, the presence, in said nucleic acid sample, of a mutation in the *ace-1* gene as defined in claim 6 or in claim 7.

10. The method as claimed in claim 9, characterized in that said detection comprises:

- amplifying a fragment of approximately 320 bp using the pair of primers SEQ ID NOs. 39 and 40,

- digesting said fragment using a suitable restriction enzyme, and

- analyzing the restriction profile obtained.

11. The method as claimed in claim 10, characterized in that said restriction enzyme is *EcoRI*.

12. A recombinant vector, characterized in that it comprises an insert selected from the group consisting of the nucleic acid molecules as claimed in any one of claims 6 to 8.

13. A cell, characterized in that it is modified with a recombinant vector as claimed in claim 12.

14. An antibody, characterized in that it is directed against the acetylcholinesterase as claimed in any one of claims 1 to 4 or the peptide as claimed in claim 5.

15. A reagent for detecting insects carrying resistance to insecticides of the organophosphorus compound and carbamate class, characterized in that it is selected from the group consisting of the nucleic

acid molecules and the fragments thereof as claimed in any one of claims 6 to 8 and the antibodies as claimed in claim 14.

16. A transgenic invertebrate animal, characterized in that it contains cells transformed with at least one nucleic acid molecule as claimed in claim 6 or claim 7.

17. A method for screening an insecticidal substance, characterized in that it comprises:

a) bringing the test substance into contact with an acetylcholinesterase as claimed in any one of claims 1 to 4, or an extract of modified cells as defined in claim 13, or a biological sample from a transgenic animal as defined in claim 16, in the presence of acetylcholine or of one of its derivatives, and

b) measuring, by any suitable means, the acetylcholinesterase activity of the mixture obtained in a), and

c) selecting the substances capable of inhibiting said activity.

18. A method for screening insecticidal substances, characterized in that it comprises:

- bringing a test substance into contact with a transgenic animal as claimed in claim 16, and
- measuring the animal's survival.

19. A reagent for screening insecticidal substances, characterized in that it is selected from the group consisting of the acetylcholinesterases as claimed in any one of claims 1 to 4, the recombinant vectors as claimed in claim 12, the modified cells as claimed in claim 13 and the transgenic animals as claimed in claim 16.

20. A detection and/or screening kit, characterized in that it includes at least one reagent as claimed in claim 15 or claim 19.

ABSTRACT

novel acetylcholinesterase gene (*ace-1*) responsible for resistance to organophosphorus and carbamates in mosquitoes, non-homologous to the *D. melanogaster* acetylcholinesterase gene (*ace-2*), products of the *ace-1* gene (cDNA, protein AchE1) and the applications thereof, particularly for the screening of novel insecticides and the genetic detection of resistance to organophosphorus and carbamates in mosquito populations.

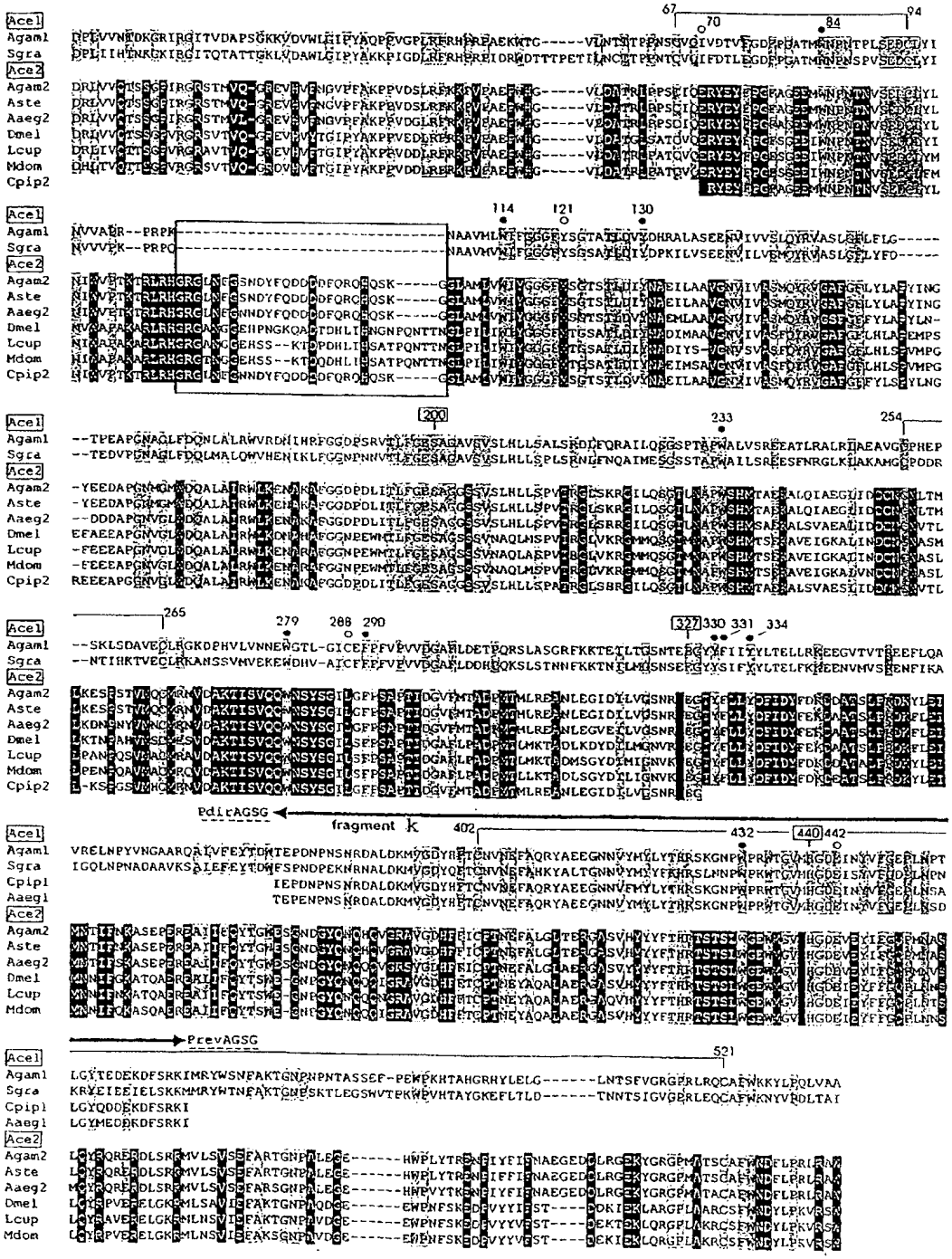
Ace1
Agam1
Sgra
Ace2
Agam2
Aste
Aaeg2
Dmel
Lcup
Mdom
Cpip2
Cpip1
Aaeg1
67
70
84
94
114
121
130
200
233
254
265
279
288
290
327
330
331
334
402
432
440
442
521
PdirAGSG
fragment k
PrevAGSG

FIGURE 1

A

```
                1                                                                  80
Ae alb    TEPENPNSNR DALDKMVGDY HFTCNVNEFA QRYAEEGNNV YMYLYTHRSK GNPWPRWTGV MHGDEINYVF GEPLNSDLGY
Ae aeg    TEPENPNSNR DALDKMVGDY HFTCNVNEFA QRYAEEGNNV YMYLYTHRSK GNPWPRWTGV MHGDEINYVF GEPLNSDLGY
An alb    TEPDNPNSNR DALDKMVGDY HFTCNVNEFA QRYAEEGNNV YMYLYTHRSK GNPWPRWTGV MHGDEINYVF GEPLNPTLGY
An gam    TEPDNPNSNR DALDKMVGDY HFTCNVNEFA QRYAEEGNNV YMYLYTHRSK GNPWPRWTGV MHGDEINYVF GEPLNPTLGY
An fun    TEPDNPNSNR DALDKMVGDY HFTCNVNEFA QRYAEEGNNV YMYLYTHRSK GNPWPRWTGV MHGDEINYVF GEPLNPSLGY
An nil    TEPDNPNSNR DALDKMVGDY HFTCNVNEFA QRYAEEGNNV YMYLYTHRSK GNPWPRWTGV MHGDEINYVF GEPLNPSLGY
An sac    TEPDNPNSNR DALDKMVGDY HFTCNVNEFA QRYAEEGNNV YMYLYTHRSK GNPWPRWTGV MHGDEINYVF GEPLNPSLGY
An pse    TEPDNPNSNR DALDKMVGDY HFTCNVNEFA QRYAEEGNNV YMYLYTHRSK GNPWPRWTGV MHGDEINYVF GEPLNPGLGY
Cx Pip    TEPDNPNSNR DALDKMVGDY HFTCNVNEFA QRYAEEGNNV YMYLYTHRSK GNPWPRWTGV MHGDEINYVF GEPLNSALGY

             81        91
Ae alb    MDDEKDFSRK I
Ae aeg    MEDEKDFSRK I
An alb    TDDEKSFSRK I
An gam    TEDEKDFSRK I
An fun    TEDEKDFSRK I
An nil    TEDEKDFSRK M
An sac    TDDEKDFSRK I
An pse    TEDEKDFSRK I
Cx Pip    QDDEKDFSRK I
```

B

```
                          •         20         •         40         •         60         •         80
Ace1-SLAB    ATCGAACCGGACAACCCGAACAGCAACCGTGACGCGCTGGACAAGATGGTCGGGGATTATCACTTCACCTGCAACGTGAA
Ace1-SR      ATCGAACCGGACAACCCGAACAGCAACCGTGACGCGCTCGACAAGATGGTCGGGGATTATCACTTCACCTGCAACGTGAA

                EcoRI     •        100         •        120         •        140         •        160
Ace1-SLAB    CGAGTTCGCCCAGCGGTACGCCGAGGAGGGCAACAACGTGTTCATGTACCTGTACACGCACAGAAGCAAAGGAAATCCCT
Ace1-SR      CGAGTTCGCCCAGCGGTACGCCGAGGAGGGCAACAATGTGTTCATGTACCTGTACACGCACAGAAGCAAAGGAAATCCCT

                          •        180         •        200         •        220         •        240
Ace1-SLAB    GGCCGAGGTGGACCGGCGTGATGCACGGCGACGAGATCAACTACGTGTTTGGCGAACCGCTGAACTCGGCCCTCGGCTAC
Ace1-SR      GGCCGAGGTGGACCGGCGTGATGCACGGCGACGAGATCAACTACGTGTTTGGCGAACCGCTGAACTCGGCCCTCGGCTAC

                                  •        260         •
Ace1-SLAB    CAGGACGACGAGAAGGACTTTAGCCGGAAAATT
Ace1-SR      CAGGACGACGAGAAGGACTTTAGCCGGAAAATT
```

C

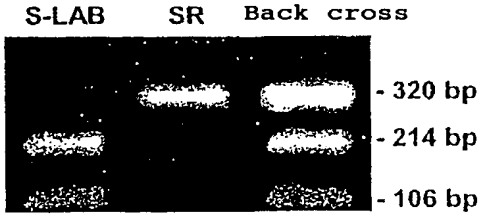

FIGURE 2

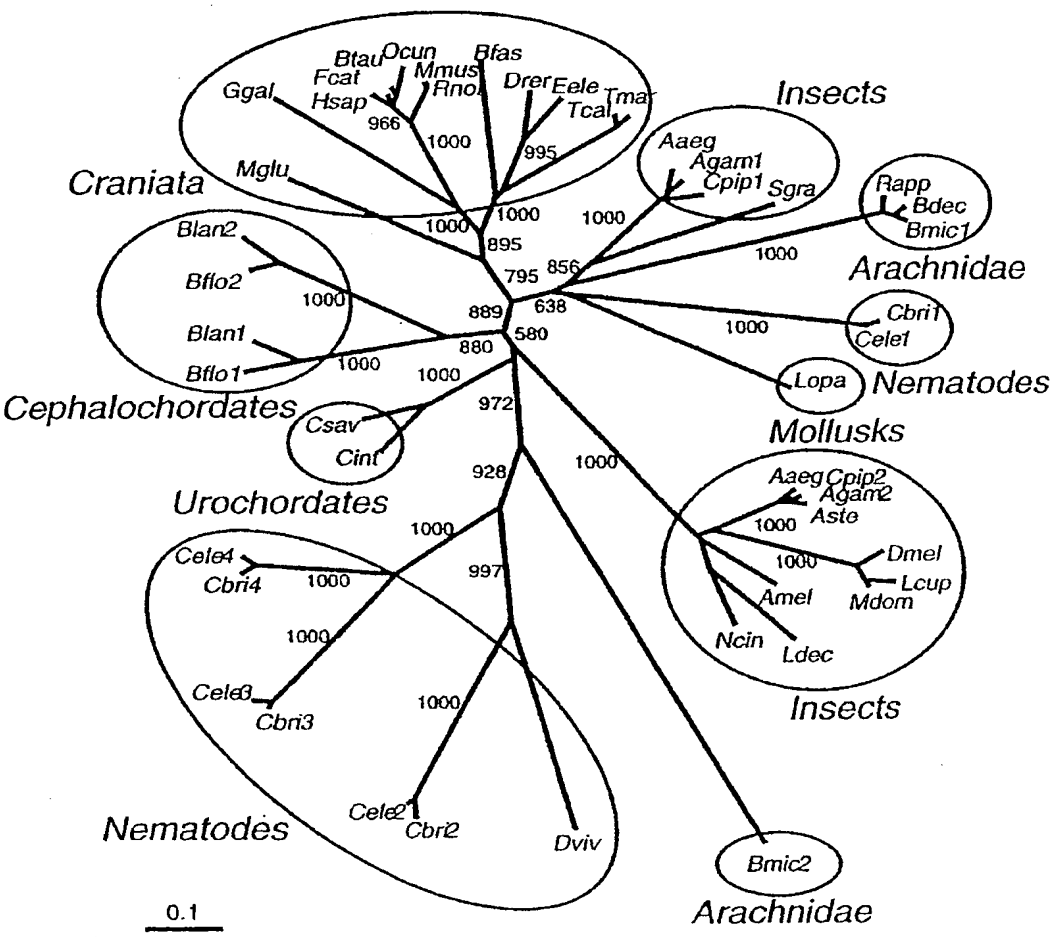
Craniata
Insects
Arachnidae
Nematodes
Mollusks
Cephalochordates
Urochordates
Nematodes
Insects
Arachnidae
Btau
Ocun
Bfas
Fcat
Mmus
Rno
Drer
Eele
Tma
Ggal
Hsap
Tcal
966
1000
995
Aaeg
Agam1
Cpip1
Sgra
Mglu
Rapp
Bdec
Bmic1
1000
1000
1000
895
795
856
1000
Blan2
Bflo2
1000
889
638
Cbri1
Cele1
1000
Blan1
880
580
Bflo1
1000
1000
Lopa
972
Csav
Cint
928
1000
Aaeg
Cpip2
Agam2
Aste
1000
Cele4
Cbri4
1000
997
1000
Dmel
Lcup
Mdom
Amel
Ncin
Ldec
Cele3
Cbri3
1000
Cele2
Cbri2
Dviv
Bmic2
0.1

FIGURE 3

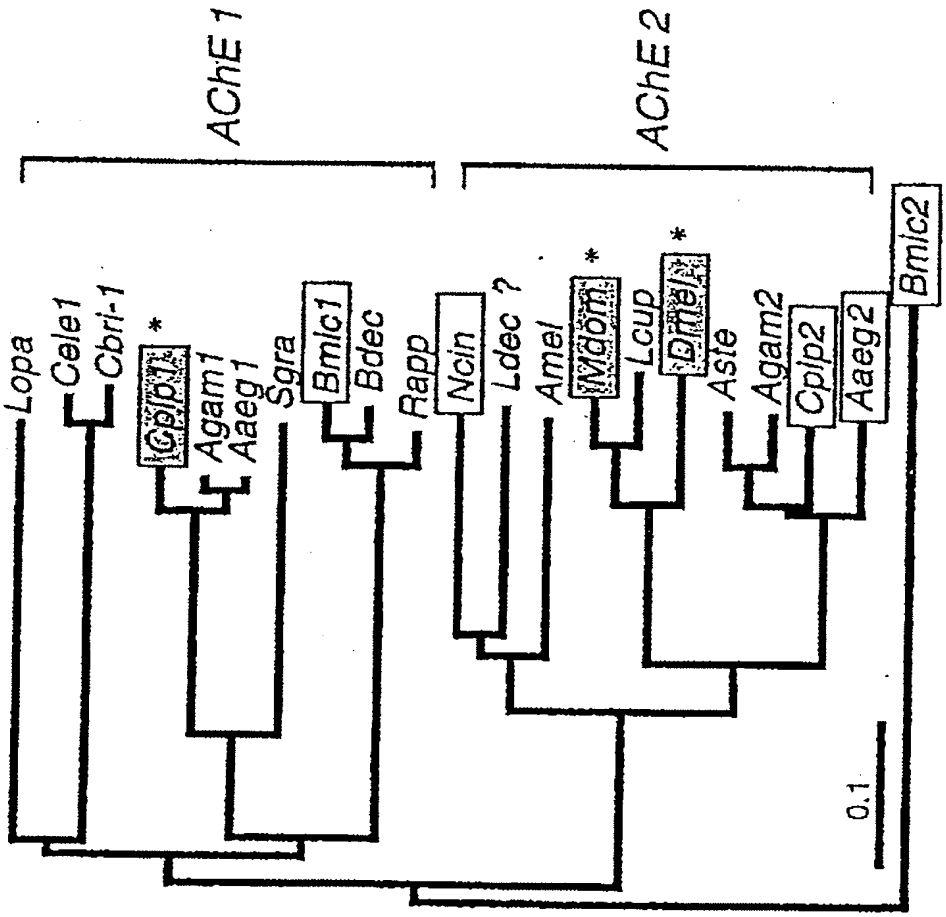
AChE 1
AChE 2
Lopa
Cele1
Cbri-1
Cplp1
Agam1
Aaeg1
Sgra
Bmic1
Bdec
Rapp
Ncin
Ldec ?
Amel
Mdom
Lcup
Dmel
Aste
Agam2
Cplp2
Aaeg2
Bmic2
*
*
0.1

FIGURE 4